Relative Risk of Alzheimer Disease and Age-at-Onset Distributions, Based on *APOE* Genotypes among Elderly African Americans, Caucasians, and Hispanics in New York City

Ming-Xin Tang,[1,2] Gladys Maestre,[2,7] Wei-Yann Tsai,[4] Xin-Hua Liu,[1] Lin Feng,[7] Wai-Yee Chung,[7] Michael Chun,[5] Peter Schofield,[5] Yaakov Stern,[1,4,5,6] Benjamin Tycko,[2,7] and Richard Mayeux[1,2,3,5,6]

[1]The Gertrude H. Sergievsky Center, [2]The Taub Center for Alzheimer's Disease Research in the City of New York, the Divisions of [3]Epidemiology and [4]Biostatistics, School of Public Health, and the Departments of [5]Neurology, [6]Psychiatry, and [7]Pathology, Columbia University College of Physicians and Surgeons, New York

Summary

Apolipoprotein-E ε4 (*APOE*-ε4) has been consistently associated with Alzheimer disease (AD) and may be responsible for an earlier age at onset. We have previously reported a diminished association between *APOE*-ε4 and AD in African Americans. Using a new method, which allows inclusion of censored information, we compared relative risks by *APOE* genotypes in an expanded collection of cases and controls from three ethnic groups in a New York community. The relative risk for AD associated with *APOE*-ε4 homozygosity was increased in all ethnic groups (African American relative risk [RR] = 3.0; 95% confidence interval [CI] = 1.5–5.9; Caucasian RR = 7.3, 95% CI = 2.5–21.6; and Hispanic RR = 2.5, 95% CI = 1.1–5.7), compared with those with *APOE*-ε3/ε3 genotypes. The risk was also increased for *APOE*-ε4 heterozygous Caucasians (RR = 2.9, 95% CI = 1.7–5.1) and Hispanics (RR = 1.6, 95% CI = 1.1–2.3), but not for African Americans (RR = 0.6, 95% CI = 0.4–0.9). The age distribution of the proportion of Caucasians and Hispanics without AD was consistently lower for *APOE*-ε4 homozygous and *APOE*-ε4 heterozygous individuals than for those with other *APOE* genotypes. In African Americans this relationship was observed only in *APOE*-ε4 homozygotes. These results confirm that *APOE* genotypes influence the RR of AD in Caucasians and Hispanics. Differences in risk among *APOE*-ε4 heterozygote African Americans suggest that other genetic or environmental factors may modify the effect of *APOE*-ε4 in some populations.

Introduction

The presence of the ε4 allele of the apolipoprotein-E gene (*APOE*) has been consistently associated with an increased risk of Alzheimer disease (AD) (Borgaonkar et al. 1993; Corder et al. 1993, 1994; Mayeux et al. 1993; Noguchi et al. 1993; Payami et al. 1993; Poirier et al. 1993; Saunders et al. 1993*a*, 1993*b*; Strittmatter et al. 1993*a*, 1993*b*; Chartier-Hardlin et al. 1994; Kuusisto et al. 1994; Liddell et al. 1994; Roses et al. 1994; Sorbi et al. 1994; Hendrie et al. 1995; Maestre et al. 1995). However, it is apparent that the *APOE*-ε4 allele alone is neither necessary nor sufficient to cause the disease, because (1) AD develops in the absence of *APOE*-ε4 (Lannfelt et al. 1994; Maestre et al. 1995) and (2) some individuals escape the disease despite having an *APOE*-ε4 allele. The presence of *APOE*-ε2 has been associated with both a decreased (Corder et al. 1994; Chartier-Hardlin et al. 1994) and an increased risk of AD (Lannfelt et al. 1994; Maestre et al. 1995; van Duijn et al. 1995).

Roses et al. (Roses et al. 1994; Roses 1995) have explained the variability by proposing that *APOE* genotypes influence the age at onset of AD. In both familial and sporadic AD, they observed an earlier age at onset among *APOE*-ε4 homozygous and *APOE*-ε4 heterozygous AD cases than among those AD cases with other *APOE* genotypes. These observations have needed confirmation, because they imply that *APOE* genotypes may predict specific age-at-onset distributions. Commercial tests based on the *APOE* allele system have become available, allowing physicians to estimate the risk of AD and the average age at onset, by *APOE* genotype.

There may be important limitations, however. The association between AD and *APOE*-ε4 is less consistent for individuals >80 years of age (Poirier et al. 1993; Roses et al. 1994), when disease risk is highest (Evans et al. 1989; Gurland et al. 1995; Hebert et al. 1995), and the association may be stronger in women than in men (Payami et al. 1994). The relationship between AD

and APOE-ε4 in ethnic groups other than Caucasians has not been thoroughly examined. For example, an association between AD and APOE-ε4 homozygosity, but not between AD and APOE-ε4 heterozygosity, was observed in African Americans in New York (Maestre et al. 1995), whereas both APOE-ε4 homozygosity and APOE-ε4 heterozygosity were associated with AD among African Americans in Indiana (Hendrie et al. 1995).

In this investigation we developed a new method to estimate relative risk (RR) of AD among individuals, with incomplete observations based on APOE genotypes in an expanded elderly group of African Americans, Hispanics, and Caucasians residing in the Washington Heights community of New York City. We also used a previously published method (Turnbull 1974) to plot the age distribution of the proportion of individuals who remained free of AD, by APOE genotype, within each ethnic group.

Subjects and Methods

Subjects and Setting

All data were derived from individuals residing in the same community. AD cases were identified as part of a comprehensive community-based registry of all dementias in northern Manhattan. We used information (clinical and billing records, the computerized daily census, monthly discharge lists, and the daily logs) from a number of sources, medical as well as nonmedical: regional hospitals (including inpatient and outpatient services), private practitioners and the single neurologist unaffiliated with the Columbia Presbyterian Medical Center, federal and state health agencies offering assistance to individuals in the area, nursing homes and other health-related facilities, an HMO (Health Insurance Plan of Greater New York), and all senior-citizen centers in the community. A list of individuals with home–health-care attendants in the community was obtained from the Human Resources Administration in New York City. All individuals were screened to identify those with and without cognitive impairment, as described elsewhere (Gurland et al. 1995). Those with, and a proportion of those without, cognitive impairment were examined for the presence or absence of AD. Controls were volunteers recruited from the same community, but most were identified in senior-citizen centers and not in medical facilities. As with AD cases, all controls at each community site were informed about a study of aging and health and were asked to participate. All controls received interviews and clinical assessments identical to those used for AD cases. The refusal rate, for both AD cases and controls identified at these centers, was <20%.

AD cases and controls were also identified from a random survey of Medicare recipients in the same geographic area (the population >65 years of age was 9,349; United States Bureau of the Census 1991). The Health Care Financing Administration (HCFA) provided access to a random sample of the names and addresses of recipients within three contiguous ZIP codes in the community. Potential subjects were sent a letter from HCFA that explained that they had been randomly selected as candidates for a study of aging. A 90-min in-person interview was conducted for those who agreed to participate, followed by a complete medical and neurological assessment by a physician. Names of AD cases and controls identified in the survey and the registry were cross-referenced to eliminate duplicates. Participation rate was 68% in the HCFA-based random sample and did not differ by ethnic group or gender.

The Columbia University institutional review board reviewed and approved this project. All individuals provided written informed consent. Both AD cases and controls were followed over a 4-year period.

Diagnosis of AD

Physicians elicited the medical history and conducted a standardized physical and neurological examination. Whenever possible, all ancillary information, including medical charts and reports of laboratory studies, were used in the evaluation. A standardized neuropsychological battery measuring performance in memory, orientation, abstract reasoning, language, and construction was administered. Independence in activities of daily living was measured to determine whether subjects met functional criteria for dementia. Except for APOE genotypes, all clinical information was reviewed at a diagnostic conference of physicians and neuropsychologists, to arrive at a consensus diagnosis based on research criteria (McKhann et al. 1984; American Psychiatric Association 1987). The development of our diagnostic methods in relation to the cultural and educational demographics of this community has been described elsewhere (Pittman et al. 1992; Stern et al. 1992). The reliability and consistency of these assessments have also been reported elsewhere (Schofield et al., in press). The majority of AD cases were alive at the time of this investigation, but data were also available from 16 AD cases with postmortem confirmation of AD that were from the same community.

Age at Onset

Although all informants for prevalent AD cases were asked about the duration of symptoms, we did not attempt to estimate the age at onset in these individuals, because of the difficulty in substantiating this type of information. To improve validity, we censored (left-censored) the age at onset in prevalent AD cases, on the

basis of methods described elsehwere (Turnbull 1974; Cupples et al. 1991; Tang et al., in press). In brief, for each prevalent case, the EM (expectation and maximization) algorithm, as implemented by Turnbull, was used to compute the expected age at onset (see appendixes). This value was less than the patient's age on entry into the study and was within an age interval based on the distribution of ages either (1) at study entry, for all younger prevalent AD cases, or (2) ages at onset, for all younger incident AD cases. In the age interval, the lowest age for Caucasians and Hispanics was 50 years, whereas in African Americans the lowest age was 59 years. We used age 50 years as the lowest age for computing the expected age at onset for all prevalent AD cases, regardless of ethnic group. For example, in a patient with AD who was 75 of age, the estimated age at onset would be 50–75 years. Several controls in the registry became demented during the 4-year follow-up. The age at onset for these AD cases was their age on the date at which dementia was first documented. For controls, we right-censored at their age at the last follow-up examination.

Ethnic Group

Ethnic group was classified by self-report using the format of the 1990 United States Bureau of the Census (1991). This allows each individual to identify himself as belonging to a particular racial or ethnic group. Each individual is then asked whether he or she is of "Hispanic" origin. Using this information, we separated subjects into three ethnic groups: African American, Hispanic, and Caucasian (non-Hispanic), according to self-report, on the basis of direct interview with the subject or, if the individual was demented, a family member.

APOE Genotyping

Genomic DNA was amplified by PCR and was subjected to CfoI restriction analysis using APOE primers and conditions modified from those described by Hixson and Vernier (1991). Genotypes were determined without knowledge of patient-control status.

Data Analysis

APOE allele frequencies were determined by counting alleles and calculating sample proportions. For comparison across ethnic groups, APOE allele frequencies were calculated by use of data from all controls and from a random 10% sample of the AD cases within each ethnic group, to create a representative population of individuals ≥ 65 years of age that would include patients with AD (Evans et al. 1989). For comparison of AD cases and controls, within and across ethnic groups, APOE allele frequencies were calculated for all subjects and were compared by χ^2 analysis (Fleiss 1981).

We used an iterative method (i.e, the EM algorithm) to obtain the nonparametric maximum-likelihood estimates of the age-at-onset distributions, by APOE genotypes, within each ethnic group, from the "self-consistent equation" (Kaplan and Meier 1958; Turnbull 1974; present study, appendix A). A parametric approach was used to develop a proportional hazards regression model to estimate the RR for AD, by APOE genotype (refer to appendix B). This analysis was similar to the method used by Cupples et al. (1989, 1991). Both univariate and multivariate RRs for AD associated with various APOE genotypes were derived from the proportional hazards model.

Results

Demographics

Data were available for 305 patients with AD and for 485 nondemented healthy elderly. Some 55% of the patients with AD were identified through the registry, compared with 45% from the survey. Some 31% of the controls were volunteers from the community sites, and 69% were identified through the survey. Patients and controls differed by gender (76% of AD cases were women, vs. 67% of controls; $\chi^2 = 6.8$, $P < .01$). In all ethnic groups combined, AD cases were older (76.4 ± 9.1 years [range 46–96 years]) than controls (72.9 ± 6.7 years [range 46–93 years]; $P < .05$). African Americans and Hispanics were generally older and less well educated than Caucasians (table 1). African-American and Hispanic AD cases also had less education than controls (7.3 ± 5.0 years for AD cases, vs. 9.0 ± 4.6 years for controls; $P < .05$).

APOE Allele Frequencies

Overall, the APOE allele frequencies differed significantly among the three ethnic groups (data not shown; $\chi = 19.7$; $P < .0005$). The main difference was in the frequency of APOE-ε4 among African Americans, which was most pronounced among individuals >72 years of age (median age), in whom the frequency of APOE-ε4 was .21, compared with .09 both in Caucasians and in Hispanics (table 2). APOE allele frequencies did not differ by gender.

Among AD cases, the distribution of APOE alleles also differed across ethnic groups ($\chi^2 = 11.9$, df 4, $P < .01$), with higher frequencies of APOE-ε4 in Caucasians (27%) than in African Americans (23%) or Hispanics (19%) and with higher frequencies of APOE-ε2 in African Americans (11%) than in the other two ethnic groups (Caucasians, 3%; and Hispanics, 7%), as noted in table 1. In Caucasians and Hispanics, but not in African Americans, the APOE-ε4 allele frequency was significantly higher among AD cases than among controls.

Table 1

Demographics and *APOE* Allele Frequency among Patients with AD and among Controls, by Ethnic Group

ETHNIC GROUP (N)	AGE (years)	EDUCATION (years)	FREQUENCY		
			APOE-ε3	*APOE*-ε4	*APOE*-ε2
African American:					
AD cases (106)	78 ± 7.6	7.4 ± 4.0	.66	.23	.11
Controls (154)	74 ± 5.8[a]	9.4 ± 3.8[a]	.73	.20	.07
Caucasian:					
AD cases (59)	71 ± 9.1	13.4 ± 4.2	.70	.27	.03
Controls (112)	72 ± 8.5	12.3 ± 4.0[a]	.84	.09	.07
Hispanic					
AD cases (140)	77 ± 8.3	4.8 ± 3.7	.74	.19	.07
Controls (219)	73 ± 6.1[a]	7.0 ± 4.3[a]	.83	.10	.07

NOTE.—Allelic frequencies differed significantly between AD cases and controls, within and across ethnic groups; for explanation, see the text.

[a] Significant difference ($P < .05$) exists within each stratum.

When the groups were subdivided into age quartiles, however, *APOE*-ε4 frequency in African-American AD cases in the youngest quartile (age ≤69 years) differed significantly from that in controls in the youngest quartile (42% in AD cases, vs. 15% in controls), because of a greater number of *APOE*-ε4 homozygotes in this age group. In the three other quartiles, African-American AD cases and controls had nearly identical *APOE*-ε4 frequencies.

RR and the Age-at-Onset Distribution

The impact of *APOE* genotypes on RR was calculated, and the cumulative age-at-onset distributions were

Table 2

***APOE* Allele Frequency, by Age and Ethnic Group**

ETHNIC GROUP	FREQUENCY (No. Counted)		
	APOE-ε3	*APOE*-ε4	*APOE*-ε2
African American:			
≤72 years of age	.72 (107)	.20 (27)	.08 (12)
>72 years of age[a]	.74 (136)	.21 (38)	.05 (10)
Caucasian:			
≤72 years of age	.84 (112)	.10 (13)	.06 (9)
>72 years of age	.85 (87)	.09 (9)	.06 (6)
Hispanic:			
≤72 years of age	.81 (193)	.11 (26)	.08 (19)
>72 years of age	.86 (195)	.09 (21)	.05 (12)

NOTE.—To compare ethnic groups across ages, we used all controls recruited and a 10% random sample of cases from each ethnic group, to form a representative sample of elderly ≥65 years of age.

[a] $\chi^2 = 19.7$, df = 4, $P < .0005$. *APOE* allele frequencies differ in African Americans relative to other ethnic groups >72 years of age, the median age of the cohort.

plotted for each ethnic group. Specific *APOE* genotypes for AD cases and controls, for each ethnic group, are shown in table 3. African Americans homozygous for the *APOE*-ε4 allele had a significantly increased risk of AD (RR = 3.0, 95% CI = 1.5–5.9) compared with African Americans with *APOE*-ε3/ε3 genotypes, but the risk associated with *APOE*-ε4 heterozygosity was significantly decreased (RR = 0.6, 95% CI = 0.4–0.9). The proportions, by age and *APOE* genotype, of individuals without AD are shown in figure 1. We isolated individuals with *APOE*-ε4 homozygosity and individuals with *APOE*-ε4 heterozygosity, to show the differences between the age distributions in the proportion remaining unaffected, compared with other *APOE* genotypes.

Caucasians, both those *APOE*-ε4 homozygous and those *APOE*-ε4 heterozygous, had a significantly increased RR for AD (ε4 homozygosity RR = 7.3, 95% CI = 2.5–21.6; and ε4 heterozygosity RR = 2.9, 95% CI = 1.7–5.1), compared with Caucasians with *APOE*-ε3/ε3 genotypes. Hispanics with one or more copies of the *APOE*-ε4 allele also had higher RRs of AD (ε4 homozygous RR = 2.5, 95% CI = 1.1–5.7; and ε4 heterozygous individuals RR = 1.6, 95% CI = 1.1–2.3), compared with Hispanics with *APOE*-ε3/ε3 genotypes. Figures 2 and 3 illustrate the age distribution of the proportion of individuals without AD, comparing those with at least one *APOE*-ε4 allele to those with other *APOE* genotypes, among Caucasians and Hispanics.

We also examined the protective effect of *APOE* genotypes with one or more ε2 alleles, again using *APOE*-ε3/ε3 genotypes as the reference. The disease risk associated with *APOE*-ε2 was lower among Caucasians (RR = 0.2, 95% CI 0.03–1.7), but not in African Americans (RR = 1.3, 95% CI = 0.8–2.3) or Hispanics (RR = 1.0,

Table 3

APOE Genotypes in AD Cases and in Controls, by Ethnic Group

	No. (Proportion) of Individuals					
Ethnic Group	ε3/ε3	ε2/ε3	ε2/ε2	ε4/ε2	ε3/ε4	ε4/ε4
African American:						
AD cases	49 (.46)	16 (.151)	3 (.028)	2 (.019)	26 (.245)	10 (.094)
Controls	77 (.50)	16 (.104)	1 (.006)	3 (.019)	55 (.357)	2 (.013)
Caucasian:						
AD cases	30 (.508)	1 (.017)	0	2 (.034)	22 (.373)	4 (.068)
Controls	78 (.696)	14 (.125)	0	1 (.009)	19 (.17)	0
Hispanic:						
AD cases	80 (.571)	12 (.086)	1 (.007)	5 (.036)	36 (.257)	6 (.043)
Controls	148 (.676)	29 (.132)	1 (.015)	0	38 (.174)	3 (.014)

95% CI = 0.5–1.7). Each analysis was repeated, with either stratification by gender or inclusion of educational level as a covariate; but the results were unchanged.

We also examined differences between the number of individuals either left- or right-censored and the number of those with complete information, by *APOE* genotype and ethnic group; there were no statistically significant differences, in the number of individuals, by *APOE* genotype, but incident AD cases were significantly more numerous among African Americans and Hispanics than among Caucasians (table 4).

Discussion

We examined the RR for AD, on the basis of *APOE* genotypes among three distinct ethnic groups residing in New York City, using a new method that included both left- and right-censored observations. Using this method, we also plotted the age distribution of the proportions of individuals within each ethnic group remaining free of AD, by *APOE* genotype. As expected, the *APOE*-ε4/ε4 genotype was associated with the highest RR for AD in all three ethnic groups, compared with individuals with the *APOE*-ε3/ε3 genotype. Among Caucasians and Hispanics, *APOE*-ε4 heterozygosity was also associated with a significantly higher RR for AD, compared with individuals with the *APOE*-ε3/ε3 genotype. However, among African Americans, *APOE*-ε4 heterozygotes had a significantly lower RR for AD, compared with individuals with the *APOE*-ε3/ε3 genotype. Stratification by gender or educational level did not change the associations between *APOE*-ε4 and AD in any ethnic group. There was only minimal evidence that AD risk was reduced for Caucasians with *APOE*-ε2 genotypes, and there was no decrease in risk for African Americans or Hispanics with these genotypes.

An association between *APOE*-ε4 and AD has been reported among African Americans in Indiana (Hendrie et al. 1995). African Americans were also included in a North Carolina study (Saunders et al. 1993a), but the specific association between *APOE*-ε4 and AD was not described. In both studies the *APOE*-ε4 frequency was lower than that reported here; in Indiana the *APOE*-ε4 allele frequency in controls was .12, whereas in North Carolina study the *APOE*-ε4 frequency was .17 overall. No association between *APOE*-ε4 and AD was found in a small study of elderly Nigerians (Osuntokun et al. 1995), where the ε4 allele frequency in the control population was .26. Differences in *APOE* allele frequencies by ethnic group are well established in adults and children (Davignon et al. 1988; Kamboh et al. 1989; Sepehrnia et al. 1989; Hallman et al. 1991; Srinivasan et al. 1993; Sandholzer et al. 1995; Zekraoui et al. 1995). Regional differences in the frequency of *APOE* alleles among African populations have also been described (Sepehrnia et al. 1989; Sandholzer et al. 1995; Zekraoui et al. 1995). Zekraoui et al. (1995) observed the *APOE*-ε4 frequency to vary from .03 among Wolof in Senegal to .38 among Tutsi in Burundi, and a similarly elevated ε4 frequency (.4) was also observed in South African bushman (Sandholzer et al. 1995). The *APOE*-ε4 allele frequency among the African Americans in the present study and others (Kamboh et al. 1989; Srinivasan et al. 1993) was 20%–26%. Thus, it is possible that the individuals included in our study were genetically distinct from those in the Indiana study. It is also important to note that the *APOE*-ε4 frequency did not decline with age, in contrast to the trend noted among Caucasians in the United States and Europe (Kervinen et al. 1994; Louhija et al. 1994; Schachter et al. 1994).

The decline in the frequency of *APOE*-ε4 with age in Caucasians has been attributed to mortality from heart disease (Eichner et al. 1993; Kervinen et al. 1994; Luc et al. 1994; Wilson et al. 1994; Stengard et al. 1995), because
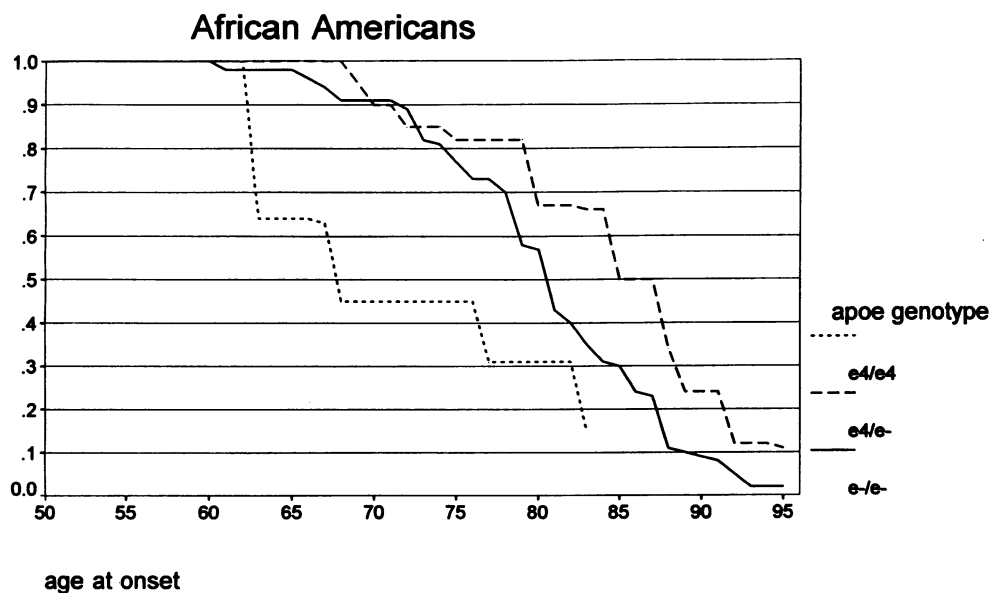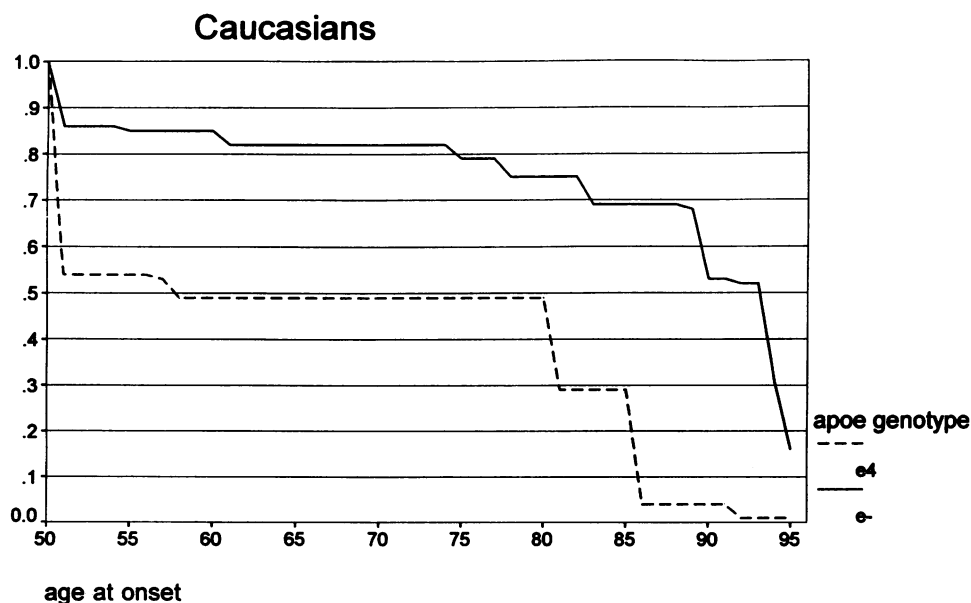
African Americans

Figure 1   Proportion of individuals without AD, among African Americans, by *APOE* genotypes. *APOE* genotypes are divided into three groups: *APOE*-ε4 homozygosity, *APOE*-ε4 heterozygosity, and all other *APOE* genotypes as the reference. We retained three groups only among African Americans, to demonstrate the different effects of *APOE* homozygosity and *APOE* heterozygosity.

variations in the *APOE* locus are associated with cholesterol and triglyceride concentrations (Davignon et al. 1988; Reilly et al. 1992; Louhija et al. 1994). For example, in young and middle-aged Finns the allele frequency of *APOE*-ε4 is ∼.2, and it decreases by almost 50% with age, because of heart disease (Kuusisto et al. 1994; Stengard et

al. 1995), but a strong association between AD and *APOE*-ε4 persists (Kuusisto et al. 1994).

In contrast, *APOE*-ε4 does not consistently predict cholesterol or triglyceride concentrations in adults or children of African origin (Sepehrnia et al. 1989; Srinivasan et al. 1993; Garjra et al. 1994; Sandhoizer et al.

Caucasians

Figure 2   Proportion of individuals without AD, among Caucasians, by *APOE* genotypes. *APOE* genotypes are divided into two groups: those with any *APOE*-ε4 and all other *APOE* genotypes as the reference. For the method used, see the text.
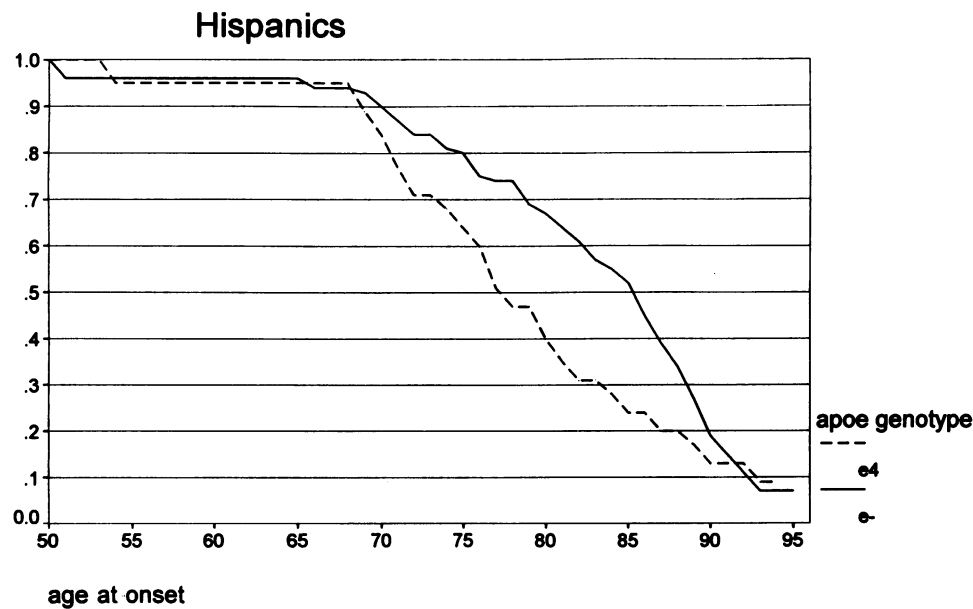
Hispanics

Figure 3   Proportion of individuals without AD, among Hispanics genotypes. *APOE* genotypes are divided into two groups: those with any *APOE*-ε4 and all other *APOE* genotypes as the reference. For the method used, see the text.

1995). Age-related changes in the population frequency of *APOE*-ε4 in African Americans have not been thoroughly investigated, but the frequency in our study remained elevated among the individuals >80 years of age. This suggests that the relationship of *APOE*-ε4 to heart disease, longevity, and AD in African Americans may not parallel the relationship in Caucasians. Alternatively, other genes or environmental factors, which are

unique to African Americans, may have a greater influence on mortality and survival.

Given the consistent association between *APOE*-ε4 heterozygosity and AD, seen both in the literature and among Caucasians and Hispanics in the present study, we considered the possibility of a systematic bias in our investigation of African Americans. The vast majority of African-American AD cases and controls were born

Table 4

Individuals with Prevalent AD, Individuals Who Are Not Demented, and Individuals with Incident AD, by *APOE* Genotype and Ethnic Group

	No. (Proportion) of Individuals					
Ethnic Group	ε3/ε3	ε2/ε3	ε2/ε2	ε4/ε2	ε3/ε4	ε4/ε4
African American:						
Prevalent-AD cases	23 (.41)	10 (.179)	2 (.036)	2 (.036)	13 (.232)	6 (.107)
Not demented	77 (.50)	16 (.104)	1 (.006)	3 (.019)	55 (.357)	2 (.013)
Incident-AD cases	26 (.52)	6 (.12)	1 (.02)	0	13 (.26)	4 (.08)
Caucasian:						
Prevalent-AD cases	23 (.46)	1 (.02)	0	2 (.04)	21 (.42)	3 (.06)
Not demented	78 (.696)	14 (.125)	0	1 (.009)	19 (.17)	0
Incident-AD cases	7 (.778)	0	0	0	1 (.11)	1 (.11)
Hispanic:						
Prevalent-AD cases	40 (.556)	6 (.083)	1 (.014)	2 (.028)	21 (.292)	2 (.028)
Not demented	148 (.676)	29 (.132)	1 (.005)	0	38 (.174)	3 (.104)
Incident-AD cases	40 (.588)	6 (.088)	0	3 (.044)	15 (.221)	4 (.059)

Note.—There was a difference in the no. of individuals with incident AD, by ethnic group ($\chi^2 = 21.1$, df = 4, $P < .001$); Caucasians had fewer individuals.

in the United States and have resided in its northeastern region for most of their adult lives. The two methods used to identify patients with AD and controls (a community-based registry and a survey of a random sample of Medicare recipients) identified a representative population, both healthy and diseased, of African Americans, Caucasians, and Hispanics residing in the community. Despite the fact that this is one of the larger studies of *APOE* and AD, the infrequency of subjects with certain *APOE* genotypes, such as ε2/ε2, makes some of our conclusions tentative. We have previously investigated the methods used to establish the diagnosis of AD both among individuals from different cultures and among individuals with lower educational attainment (Pittman et al. 1992; Stern et al. 1992). Our methods appear to be as free of bias as is possible (Gurland et al. 1995). Two African Americans were among the postmortem AD cases included in this investigation; a single patient was *APOE*-ε4/ε4, and another was *APOE*-ε2/ε3. In the data analyses, we adjusted for gender and education as potential confounders, but this did not change our primary observations. Further, we have been unable to attribute the lower AD risk associated with *APOE*-ε4 heterozygosity to comorbidity from heart disease (Chun et al. 1995).

For the estimation of RR, more complete information was available from African Americans and Hispanics than was available from Caucasians. Although the maximum-likelihood method does not depend on the presence of uncensored information, the estimates of RR would have improved if we had had more complete information.

Therefore, we consider it possible that some as yet unidentified genetic or environmental factor reduces the risk of AD among *APOE*-ε4 heterozygous African Americans. Alternatively, genetic variations in other proteins that interact with apolipoprotein E directly or indirectly at the cellular level might account for the observations in the present study. Because *APOE*-ε4 heterozygosity (primarily *APOE*-ε4/ε3) occurs in ~25%–30% of most populations in the United States and Europe (Davignon et al. 1988), understanding this protective effect, whether genetic or environmental, would seem an important point of investigation.

Acknowledgments

Appendix A

Estimating the Survival Distribution of Age at Onset

Let X be the age at onset of AD, and let

$$Y = \begin{cases} \text{age at first visit,} & \text{if it was prevalent AD ;} \\ \text{age at last visit,} & \text{if it was not AD ;} \\ \text{age at onset,} & \text{if it was newly diagnosed} \\ & \text{or incident AD, where age} \\ & \text{at onset is known .} \end{cases}$$

We can express the available information on X by a pair of random variables (Y, Δ), where

$$\Delta = \begin{cases} 0, & \text{if } X < Y \text{ (left-censored) ;} \\ 1, & \text{if } X > Y \text{ (right-censored) ;} \\ 2, & \text{if } X = Y \text{ (uncensored) .} \end{cases}$$

Our goal is to estimate the survival distribution function $S(x) = 1 - F(x) = P(X > x)$ of X from a sample of n independent pairs (y_i, δ_i), $i = 1, \cdots, n$.

Turnbull (1974) proved that the nonparametric maximum-likelihood estimate (NMLE) $\hat{S}$ satisfies the following "empirical self-consistent equation":

$$\hat{S}(x) = S_0^e(x) + S_1^e(x) + S_2^e(x)$$
$$- \int_0^x \frac{\hat{S}(x)}{\hat{S}(t)} \, dS_1^e(t) + \int_x^\infty \frac{1 - \hat{S}(x)}{1 - \hat{S}(t)} \, dS_0^e(t) \, ,$$

where

$$S_i^e(x) = \frac{1}{n} \sum_{j=1}^n \bigg/ (y_j > x, \, \delta_j = i), \quad i = 0, 1, 2 \, ,$$

and the indicator function $I(\bullet)$ is defined by

$$I(y_j > x, \, \delta_j = i) = \begin{cases} 0, & \text{if } y_j \leq x \text{ or } \delta_j \neq i \text{ ;} \\ 1, & \text{if } y_j > x \text{ and } \delta_j = i \, . \end{cases}$$

Therefore the NMLE $\hat{S}$ is also called a "self-consistent estimator." We can use the EM algorithm to obtain the NMLE $\hat{S}$.

Appendix B

Regression Model—Parametric Approach

Regression models proposed in survival analysis often involve the assumption of proportional hazard functions; that is, $h(x,z) = r(\beta'z)h_0(x\ )$, where $h(x,z)$ is the hazard function for a subject with a p vector measure of covariates z. $r(\beta'z)$ expresses the relationship between z and regression parameter β and can be interpreted as RR. $h_0(x)$ is the hazard function of the underlying failure-time distribution. In some AD cases $h_0(x)$ has simple a form, $h_0(x) = h$ (constant) if the underlying distribution is exponential, whereas $h_0(x) = \lambda x^\gamma$ for Weibull distribution (Lawless 1982, p. 16). The most commonly used proportion function is $r(\beta'z) = \exp(\beta'z)$. The assumption can be empirically checked by using computer-graphic displays.

The parameters in the regression model are composed of two parts; one is the parameter β involved in the proportion function, and the other one, θ, characterizes the underlying distribution. The model parameter $\eta = (\theta,\beta)$ can be estimated by the maximum-likelihood method.

If it is assumed that there are no uncensored AD cases, the likelihood function can be written as

$$L(\theta, \beta) = \prod_{i=1}^{n} S(Y_i, \theta, \beta)^{\delta_i}[1 - S(Y_i, \theta, \beta)]^{1-\delta_i} .$$

With the proportional hazard model, the score function, which is the first derivative of the log likelihood, is

$$Q_1(\theta, \beta) = \frac{\partial \log L(\theta, \beta)}{\partial \theta}$$

$$= \sum_{i=1}^{n} \frac{\delta_i - S(y_i; \theta, \beta)}{1 - S(y_i; \theta, \beta)} e^{\beta'x_i} \frac{\partial \log S_0(y_i; \theta)}{\partial \theta} ;$$

$$Q_2(\theta, \beta) = \frac{\partial \log L(\theta, \beta)}{\partial \beta}$$

$$= \sum_{i=1}^{n} \frac{\delta_i - S(y_i; \theta, \beta)}{1 - S(y_i; \theta, \beta)} e^{\beta'z_i} \log S_0(y_i; \theta)z_i .$$

where S_0 is the underlying survival function. The maximum-likelihood estimate η can be obtained by numerically solving the equation $Q = 0$. The asymptotic vari-

ance of η is estimated by the inverse of the observed information matrix of second partials of the log likelihood for Newton's method (Isaacson and Keller 1966, p. 97). Note that, when X has Weibull or exponential distribution (a special case of $\gamma = 1$), the proportion hazard model is $h(x,z) = \lambda\,x^\gamma\,e^{\beta'z}$. The score function can be simplified as

$$Q(\eta) = \sum_{i=1}^{n} \frac{\delta_i - S(y_i; \eta)}{1 - S(y_i; \eta)} e^{\eta'\xi_i}\xi_i ,$$

where

$$S(y_i; \eta) = \exp[-\exp(\eta'z_i)],$$

$$\eta = [\log(\lambda), \gamma, \beta), \text{ and } \xi_i' = (1, \log y_i, z_i') .$$

The stated method can be easily modified for the data that include the uncensored AD cases.

References

American Psychiatric Association (1987) Diagnostic and statistical manual of mental disorders, 3d ed rev. American Psychiatric Association, Washington, DC, pp 205–224

Borgaonkar DS, Schmidt LC, Martin SE, Kanzer MD, Edelsohn L, Growdon J, Farrer LA (1993) Linkage of late-onset Alzheimer's disease with apolipoprotein-E type 4 on chromosome 19. Lancet 342:625

Chartier-Hardlin M-C, Parfitt M, Legarin S, Perez-Tur J, Brousseau T, Evans A, Ber C, et al (1994) ApolipoproteinE, ε4 allele as a major risk factor for sporadic early and late-onset forms of Alzheimer's disease: analysis of the 19q13.2 chromosomal region. Hum Mol Genet 3:569–574

Chun M, Liu X, Maestre G, Tycko B, Mayeux R (1995) Cardiovascular disease and the association between apolipoprotein E and Alzheimer's disease among African-Americans, Caucasians and Latinos. Neurology 45, Suppl 4:A373

Corder EH, Saunders AM, Risch NJ, Strittmatter WJ, Schmechel DE, Gaskel PC, Rimmler JB, et al (1994) Apolipoprotein E type 2 allele decreases the risk for late onset Alzheimer's disease. Nat Genet 7:180–184

Corder EH, Saunders AM, Strittmatter WJ, Schmechel DE, Gaskel PC, Small GW, Roses AD, et al (1993) Gene dose of apolipoprotein-E type 4 allele and the risk of Alzheimer's disease in late onset families. Science 261:921–923

Cupples LA, Risch N, Farrer LA, Myers RH (1991) Estimation of age at onset with missing information of onset. Am J Hum Genet 49:76–87

Cupples LA, Terrrin NC, Myers RH, D'Agostino RB (1989) Using survival methods to estimate age at onset distribution for genetic disease with an application to Huntington's disease. Genet Epidemiol 6:361–371

Davignon J, Gregg RE, Sing CF (1988) Apolipoprotein E polymorphism and atherosclerosis. Arteriosclerosis 8:1–21

Eichner JE, Kuller LH, Orchard TJ, Grandits GA, McCallum

LM, Ferrell RE, Neaton JD (1993) Relation of apolipoprotein E phenotype to myocardial infarction and mortality from coronary artery disease. Am J Cardiol 71:160–165

Evans DA, Funkenstein HH, Albert MS, Scherr PA, Cook NR, Chown MJ, Hebert LE, et al (1989) Prevalence of Alzheimer's disease in a community population of older persons: higher than previously reported. JAMA 262:2551–2556

Fleiss JL (1981) Combining evidence from fourfold tables. In: Statistical methods for rates and proportions, 2d ed. John Wiley & Sons, New York, pp 160–187

Garjra B, Candlish JK, Saha N, Mak JW, Tay JS (1994) Effect of apolipoprotein E variants on plasma lipids and apolipoproteins in the Orang Asli ('aborigines') of Malaysia. Hum Hered 44:209–213

Gurland B, Wilder D, Cross P, Lantigua R, Teresi J, Bolivar M, Barret V, et al (1995) Relative rates of dementia by multiple case definitions, over two prevalence periods, in three cultural groups. Am J Geriatr Psychiatry 3:6–12

Hallman DM, Boerwinkle E, Saha N, Sandholzer C, Menzel HJ, Császár A, Utermann G (1991) The apolipoprotein E polymorphism: a comparison of allele frequencies and effects in nine populations. Am J Hum Genet 49:338–349

Hendrie HC, Hall KS, Hui S, Unverzagt FW, Yu CE, Lahiri DK, Sahota A, et al (1995) Apolipoprotein-E genotypes and Alzheimer's disease in a community study of elderly African-Americans. Ann Neurol 37:118–120

Hebert LE, Scherr PA, Beckett LA, Albert MS, Pilgram DM, Chown MJ, Funkenstein HH, et al (1995) Age-specific incidence of Alzheimer's disease in a community population. JAMA 273:1354–1359

Hixson J, Vernier D (1991) Restriction isotyping of human apolipoprotein E by gene amplification and cleavage with HhAI. J Lipid Res 31:545–548

Isaacson E, Keller HB (1966) Analysis of numerical methods. John Wiley & Sons, New York

Kamboh MI, Sepehrnia B, Ferrell RE (1989) Genetic studies of human apolipoproteins. VI. Common polymorphism of apolipoprotein E in blacks. Dis Markers 7:49–55

Kaplan EL, Meier P (1958) Non-parametric estimation from incomplete observations. J Am Stat Assoc 53:457–481

Kervinen K, Savolainen MJ, Salokannel J, Hynninen A, Heikkinen J, Ehnholm C, Koistinen MJ, et al (1994) Apolipoprotein E and B polymorphics—longevity factors assessed in nonagenarians. Atherosclerosis 105:89–95

Kuusisto J, Koivisto K, Kervinen K, Mykkanen L, Helkala EL, Vanhanen M, Hannien T, et al (1994) Association of apolipoprotein-E phenotypes with late onset Alzheimer's disease: population based study. Br Med J 309:636–638

Lannfelt L, Lilius L, Nastase M, Viitanen M, Fratiglioni L, Eggersten G, Bergland L, et al (1994) Lack of association between apolipoprotein E allele epsilon 4 and sporadic Alzheimer's disease. Neurosci Lett 169:175–178

Lawless JF (1982) Statistical models and methods for lifetime data. John Wiley & Sons, New York

Liddell M, Williams J, Bayer A, Kaiser F, Owen M (1994) Confirmation of association between the e4 allele of apolipoprotein E and Alzheimer's disease. J Med Genetics 31:197–200

Louhija J, Miettinen HE, Kontula K, Tikkanen MJ, Miettinen

TA, Tilvis RS (1994) Aging and genetic variation of plasma apolipoproteins: relative loss of the apolipoprotein E4 phenotype in centenarians. Arterioscler Thromb 14:1084–1089

Luc G, Bard JM, Arveiler D, Evans A, Cambou JP, Bingham A, Amouyel P, et al (1994) Impact of apolipoprotein E polymorphism on lipoproteins and risk of myocardial infarction: the ECTIM study. Arterioscler Thromb 14:1412–1419

Maestre G, Ottman R, Stern Y, Gurland B, Chun M, Tang M-X, Shelanski M, et al (1995) Apolipoprotein-E and Alzheimer's disease: ethnic variation in genotypic risks. Ann Neurol 37:254–259

Mayeux R, Stern Y, Ottman R, Tatemichi TK, Tang M-X, Maestre G, Ngai C, et al (1993) The apolipoprotein ε4 allele in patients with Alzheimer's disease. Ann Neurol 34:752–754

McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM (1984) Clinical diagnosis of Alzheimer's disease: report of NINCDS-ADRDA work group under the auspices of the Department of Health and Human Services Task Force on Alzheimer's Disease. Neurology 34:939–944

Noguchi S, Murukami K, Yamada N (1993) Apolipoprotein E genotype and Alzheimer's disease. Lancet 342:737

Osuntokun BO, Sahota A, Ogunniyi AO, Gureje O, Baiewyu O, Adeyinka A, Oluwole SO, et al (1995) Lack of an association between apolipoprotein E ε4 and Alzheimer's disease in elderly Nigerians. Ann Neurol 38:463–465

Payami H, Kaye J, Heston LL, Bird T, Schellenberg GD (1993) Apolipo-protein-E genotype and Alzheimer's disease. Lancet 342:738

Payami H, Montee KR, Kaye J, Wijsman EM, Bird T, Yu C-E, Heston LL, et al (1994) Apolipoprotein-E genotype and Alzheimer's disease. JAMA 271:1316–1317

Pittman J, Andrews H, Tatemichi T, Link B, Struening E, Stern Y, Mayeux R (1992) Diagnosis of dementia in a heterogenous population: a comparison of paradigm-based diagnosis and physician's diagnosis. Arch Neurol 49:461–467

Poirier J, Davignon J, Bouthillier D, Kogan S, Bertrand P, Gauthier S (1993) Apolipoprotein E polymorphism and Alzheimer's disease. Lancet 342:697–699

Reilly Sl, Ferrell RE, Kottke BA, Sing CF (1992) The gender-specific apolipoprotein E genotype influence on the distribution of plasma lipids and apolipoproteins in the population of Rochester, Minnesota. II. Regression relationships with concomitants. Am J Hum Genet 51:1311–1324

Roses AD (1995) Apolipoprotein E genotyping in the differential diagnosis, not prediction of Alzheimer's disease. Ann Neurol 38:6–14

Roses AD, Strittmatter WJ, Pericak-Vance MA, Corder EH, Saunders AM, Schmechel DE (1994) Clinical application of Apolipoprotein E genotyping to Alzheimer's disease. Lancet 343:1564–1565

Sandholzer C, Delport R, Vermaak H, Utermann G (1995) High frequency of the apo epsilon 4 allele in Khoi San from South Africa. Hum Genet 95:46–48

Saunders AM, Schmader K, Brietner JCS, Benson MD, Brown WT, Goldfarb L, Goldgaber D, et al (1993*a*) Apolipoprotein-E ε4 allele distributions in late-onset Alzheimer's disease and in other amyloid-forming diseases. Lancet 342:710–711

Saunders AM, Strittmatter WJ, Schmechel D, St George-Hyslop PH, Pericak-Vance MA, Joo SH, Rosi BL, et al (1993*b*) Association of apolipoprotein E allele e4 with late-onset familial and sporadic Alzheimer's disease. Neurology 43:1467–1472

Schachter F, Faure-Delanef L, Guenot F, Rouger H, Froguel P, Lesueur-Ginot L, Cohen D (1994) Genetic associations with human longevity at the *APOE* and *ACE* loci. Nat Genet 6:29–32

Schofield PW, Tang M-X, Marder K, Bell K, Dooneief G, Lantigua R, Wilder D, et al. Consistency of clinical diagnosis in a community-based longitudinal study of dementia and Alzheimer's disease. Neurology (in press)

Sepehrnia B, Kamboh MI, Adams-Campbell LL, Bunker CH, Nwankwo M, Majumder PP, Ferrell RE (1989) Genetic studies of human apolipoproteins. X. The effect of the apolipoprotein E polymorphism on quantitative levels of lipoproteins in Nigerian blacks. Am J Hum Genet 45:586–591

Sorbi S, Nacmias B, Forleo P, Latorraca S, Gobbini I, Bracco L, Piacentini S, et al (1994) ApoE allele frequencies in Italian sporadic and familial Alzheimer's disease. Neurosci Lett 177:100–102

Srinivasan SR, Ehnholm C, Wattingney W, Berenson GS (1993) Apolipoprotein E polymorphism and its association with serum lipoprotein concentrations in black versus white children: the Bogalusa Heart Study. Metabolism 42:381–386

Stengard JH, Zerba KE, Pekkanen J, Ehnholm C, Nissinen A, Sing CF (1995) Apolipoprotein E polymorphism predicts death from coronary heart disease in a longitudinal study of elderly Finnish men. Circulation 91:265–269

Stern Y, Andrews H, Pittman J, Sano M, Tatemichi T, Lantigua R, Mayeux R (1992) Diagnosis of dementia in a heterogenous population: development of a neuropsychological paradigm and quantified correction for education. Arch Neurol 49:453–460

Strittmatter WJ, Saunders AM, Schmechel D, Pericak-Vance M, Enghild J, Salvesen GS, Roses AD (1993*a*) Apolipoprotein E: high affinity binding to beta-amyloid and increased frequency of type 4 allele in late-onset familial Alzheimer's disease. Proc Natl Acad Sci USA 90:1977–1981

Strittmatter WJ, Weisgraber KH, Huang D, Dong LM, Salvesen GS, Pericak-Vance M, Schmechel D, et al (1993*b*) Binding of human apolipoprotein E to β-A4 peptide: isoform-specific effects and implications for late-onset Alzheimer's disease. Proc Natl Acad Sci USA 90:8098–8102

Tang M-X, Tsai W-Y, Marder K, Mayeux R. Linear rank tests for doubly censored data. Stat Med (in press)

Turnbull BW (1974) Nonparametric estimation of survivorship function with doubly censored data. J Am Stat Assoc 69:169–174

United States Bureau of the Census (1991) Census of Population and Housing 1990. Summary tape file 1, technical documentation prepared by Bureau of the Census. Bureau of the Census, Washington, DC

van Duijn CM, de Knijff P, Wehnert A, De Voecht J, Bronzova JB, Havekes LM, Hofman A, et al (1995) Reduced survival of patients with early-onset Alzheimer's disease that carry the apolipoprotein E2 allele. Ann Neurol 37:605–610

Wilson PW, Myers RH, Larson MG, Ordovas JM, Wolf PA, Schaefer EJ (1994) Apolipoprotein E alleles, dyslipidemia, and coronary heart disease: The Framingham Offspring Study. JAMA 272:1666–1671

Zekraoui L, Lucotte G, Raisonnier A (1995) Apolipoprotein E genotypes in Africa: allelic frequencies and ethnic groups. Paper presented at the Ibsen Foundation Meeting, Paris, May 29